

# 2D-spectral estimation based on DCT and modified magnitude group delay

P. Sandeep · B. K. Shreyamsha Kumar ·  
S. V. Narasimhan

Received: 12 May 2011 / Revised: 21 December 2011 / Accepted: 22 December 2011  
© Springer-Verlag London Limited 2012

**Abstract** This paper proposes two new 2D-spectral estimation methods. The 2D-modified magnitude group delay (MMGD) is applied to 2D-discrete Fourier transform (2D-DFT) for the first and to the analytic 2D-discrete Cosine transform for the second. The analytic 2D-DCT preserves the desirable properties of the DCT (like, improved frequency resolution, leakage and detectability) and is realized by a 2D-discrete cosine transform (2D-DCT) and its Hilbert transform. The 2D-MMGD is an extension from 1D to 2D, and it reduces the variance preserving the original frequency resolution of 2D-DFT or 2D-analytic DCT, depending upon to which is applied. The first and the second methods are referred to as DFT-MMGD and DCT-MMGD, respectively. The proposed methods are applied to 2D sinusoids and 2D AR process, associated with Gaussian white noise. The performance of the DCT-MMGD is found to be superior to that of DFT-MMGD in terms of variance, frequency resolution and detectability. The performance of DFT-MMGD and DCT-MMGD is better than that of 2D-LP method even when the signal to noise ratio is low.

**Keywords** 2D-spectral estimation · 2D-discrete cosine transform · 2D-analytic DCT · Modified magnitude group delay · 2D-Hilbert transform

## 1 Introduction

The 2D-spectral estimation plays an important role in many scientific and engineering applications like array processing in communication and geophysics for direction of arrival estimation, synthetic aperture radar image formation, magnetic resonance spectroscopy, optics and fluid turbulence analysis using particle image velocity data. In practice, the size of the 2D data available may be generally small compared to that of 1D data due to limited number of sensors that can be used in a specific scenario. Similar to 1D spectral estimation, the 2D-spectral estimation based on 2D periodogram also suffers from large variance, and for a given size of 2D data, the variance reduction can be achieved by using smoothing windows but only at the expense of frequency resolution [1].

The spectral estimation based on the modified magnitude group delay (MMGD) is very effective for 1D data as it provides significant reduction in variance preserving the frequency resolution of the rectangular window [2,3]. The white noise can be a driving input for a system whose output can be treated as the signal. Further, a white noise (different) can get added at the output of a system as an observation noise. Also, irrespective of the origin of the signal, it can be associated with a white noise additively. The Z-transform of a white noise sequence has its most of the roots (zeros) close to the unit circle [3]. The truncation of a signal also results in zeros close to the unit circle. The zeros close to the unit circle result in ripples

---

S. V. Narasimhan (✉)  
Digital Signal Processing Systems Group, Aerospace Electronics  
and Systems Division, National Aerospace Laboratories,  
Bangalore 560 017, India  
e-mail: narasim@nal.res.in

S. V. Narasimhan  
Council of Scientific and Industrial Research, New Delhi, India

P. Sandeep · B. K. Shreyamsha Kumar  
Department of Electronics and Communication Engineering,  
National Institute of Technology Karnataka,  
Srinivasanagar PO, Surathkal 575 025, Karnataka, India

in the spectrum, which are the source of variance of the spectrum estimate. Thus, the driving noise of a system whose output is the signal, the associated noise with the signal and the truncation effects introduce zeros close to the unit circle, and these zeros significantly contribute to the variance. In the windowed periodogram based on discrete Fourier transform (DFT), the window converts the signal into /toward a minimum phase. This is equivalent to the decreased frequency resolution due to convolution between the Fourier transform of the original signal and that of the smoothing window, whose main lobe width is broader than that for a rectangular window. The process of windowing thus reduces the variance as it smoothes the spectral ripples / fine structure which is due to the zeros close to unit circle. This spectral smoothing is equivalent to pulling the zeros close to the unit circle towards the origin in the Z-plane. However, as this windowing process does not distinguish between the signal roots (signal poles and zeros which are away from the unit circle) and from the zeros close to unit circle (due to noise and truncation effect), the signal poles are also pulled towards the origin in the Z-plane, and this results in a poor frequency resolution of the spectral peaks. But the MMGD *removes only those zeros close to unit circle without disturbing the signal poles* and hence reduces the variance preserving the original frequency resolution [3].

The DCT due to its symmetrical extension has the improved frequency resolution, leakage and detectability. For a 1D signal, these desirable properties of the DCT have been exploited along with the MMGD in achieving a spectral estimate with a better: variance, frequency resolution and detectability [4]. This was possible due to derivation of an *analytic DCT* which has the desirable properties of the DCT. It is of interest to achieve the above said advantages to 2D signals by extending the 1D approach to a 2D-scenario.

This paper proposes two new 2D-spectral estimation methods where the 2D-MMGD is applied to 2D-DFT (DFT-MMGD) for the first and to the analytic 2D-DCT (DCT-MMGD) for the second. The new analytic 2D-DCT is derived from the 2D-DCT and its Hilbert transform. The new 2D-MMGD is an extension of MMGD from 1D to 2D signals. In the DCT-MMGD, the 2D-DCT enables reduction in truncation effects, leakage and improved frequency resolution, and the MMGD reduces the variance preserving the frequency resolution of the analytic DCT. It has been found that compared to the DFT-MMGD spectral estimation performance, the DCT-MMGD has improved variance, frequency resolution and detectability when applied to 2D sinusoids plus noise and 2D AR process associated with noise. In the presence of noise, the proposed methods have better frequency resolution and detectability than those of the 2D-LP spectral estimation [5].

## 2 Background

In this section, the background subject material required for the proposed method of 2D-spectral estimation, viz., the 2D-discrete cosine transform and the 1D group delay function, and the modified magnitude group delay (modified version of group delay) and 2D-group delay functions, will be briefly considered.

### 2.1 2D-discrete cosine transform (2D-DCT)

The 1D-DFT due to its abrupt truncation of the data (rectangular windowing) suffers from leakage/large variance, limited frequency resolution and poor detectability. But the 1D-DCT due to its symmetrical extension of the given data segment has the improved frequency resolution, leakage and detectability. Also for a 1D-signal, these desirable properties of the DCT along with MMGD [4] have been exploited in achieving a spectral estimate with a better: variance, frequency resolution and detectability.

Here, the 2D-DCT will be considered bringing out its desirable properties that are useful for a spectral estimator.

Let  $x(n_1, n_2)$  be a 2D sequence of size  $N_1 \times N_2$  and let the signal be zero outside the intervals  $0 \leq n_1 \leq N_1 - 1$  and  $0 \leq n_2 \leq N_2 - 1$ . Now derive  $y(n_1, n_2)$  from  $x(n_1, n_2)$  [1] as,

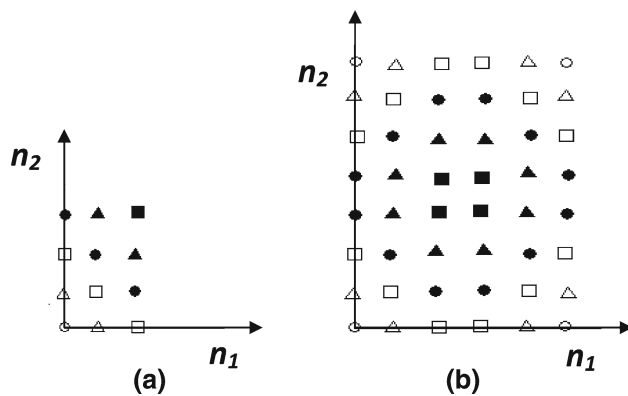
$$y(n_1, n_2) = x(n_1, n_2) + x(2N_1 - 1 - n_1, n_2) \\ + x(n_1, 2N_2 - 1 - n_2) \\ + x(2N_1 - 1 - n_1, 2N_2 - 1 - n_2) \quad (1a)$$

Let  $Y(k_1, k_2) = \text{DFT}[y(n_1, n_2)]$ , then the  $N_1 \times N_2$  DCT of  $x(n_1, n_2)$  is obtained by,

$$C_x(k_1, k_2) = \begin{cases} W_{2N_1}^{k_1/2} W_{2N_2}^{k_2/2} Y(k_1, k_2) & 0 \leq k_1 \leq N_1 - 1 \\ & 0 \leq k_2 \leq N_2 - 1 \\ 0 & \text{Otherwise} \end{cases} \quad (1b)$$

where,  $W_{2N_1} = e^{-j2\pi/2N_1}$ ,  $W_{2N_2} = e^{-j2\pi/2N_2}$   
 $y(n_1, n_2)$  is obtained by symmetrically extending  $x(n_1, n_2)$  about the origin to the second, third and fourth quadrants. This is illustrated in Fig. 1a, b for a  $3 \times 4$ -point sequence. The DFT of  $y(n_1, n_2)$  will extend this periodically.

The abruptness at the period boundaries of  $x(n_1, n_2)$  is removed for  $y(n_1, n_2)$ . Hence, the unwanted high-frequency components present in the DFT are reduced in the case of the DCT. Further, due to the smooth transition from one period to another, it is as if there is no window being applied and hence no sidelobes to cause leakage which effectively results in less spectral leakage in the DCT. Hence, the reduced truncation effects and leakage will result in less variance and less magnitude bias. The reduced magnitude bias leads to better detectability. Further, the symmetrical extension of the data



**Fig. 1** **a**  $x(n_1, n_2)$ , **b**  $y(n_1, n_2)$

will increase the frequency resolution, which may enable picking up of finer spectral details but may lead to increased variance.

## 2.2 1D-group delay function (GD), 1D-modified magnitude group delay function (MMGD) and 2D-group delay functions

For the proposed DCT-based 2D-spectral estimator, as modified group delay has to be applied, the concept of group delay function and its modified version for the 1D case are considered. Also, as the details of 2D group delay function are essential in developing the proposed spectral estimator, are covered.

### 2.2.1 1D-signal group delay function

For a minimum phase real system/signal with frequency response  $X(\omega)$ ,

$$\ln X(\omega) = \sum_{n=0}^{\infty} c(n) e^{-j\omega n} = \sum_{n=0}^{\infty} c(n) \cos(\omega n) - j \sum_{n=0}^{\infty} c(n) \sin(\omega n)$$

Also

$$\ln X(\omega) = \ln |X(\omega)| e^{j\phi(\omega)} = \ln |X(\omega)| + j\phi(\omega)$$

Therefore,

$$\ln |X(\omega)| = \sum_{n=0}^{\infty} c(n) \cos(\omega n) \quad (2a)$$

$$\theta(\omega) = \phi(\omega) + 2\pi I(\omega) = - \sum_{n=0}^{\infty} c(n) \sin(\omega n)$$

$$\theta(\omega) = - \sum_{n=0}^{\infty} c(n) \sin(\omega n) \quad (2b)$$

Further,  $\phi(\omega)$  is the principal value of the unwrapped phase  $\theta(\omega)$ , and  $I(\omega)$  is an integer-valued function required to make  $\phi(\omega)$  a continuous function of  $\omega$ . The group delay /phase group is given by

$$\tau_p(\omega) = - \frac{\partial \theta(\omega)}{\partial \omega} \quad (2c)$$

and hence the PGD  $\tau_p(\omega)$  is given by

$$\tau_p(\omega) = \sum_{n=0}^{\infty} n c(n) \cos(\omega n)$$

In Eqs. (2a) and (2b), the spectral magnitude and phase are related through the cepstral sequence  $c(n)$  and hence the group delay  $\tau_p(\omega)$  can be derived from the spectral magnitude as  $c(n)$  can be obtained using Eq. (2a). Hence, this group delay  $\tau_p(\omega)$  for a minimum phase system/signal or derived from spectral magnitude of a general system/signal is known as *magnitude group delay* (MGD) [3] and is denoted by  $\tau_m(\omega)$  which is given by

$$\tau_m(\omega) = \sum_{n=0}^{\infty} n c(n) \cos(\omega n) \quad (2d)$$

Given the spectral magnitude (Eq. (2a)), its equivalent minimum phase can be derived using Eq. (2b), and the resulting signal with this phase and the given magnitude is known as equivalent minimum phase signal derived from magnitude.

### 2.2.2 1D-modified magnitude group delay (MMGD)

The fine structure, viz., the ripples in the magnitude spectrum or the spikes in the group delay result in zeros close to the unit circle in the Z-plane and contribute significantly to the variance of the spectral estimate. Thus, the spectral ripples or variance is due to the zeros or the numerator of a transfer function. On the other hand, the signal or system spectral peaks are due to the denominator. *By normal smoothing using windows, the ripple effect or variance can be reduced but only at the cost of frequency resolution.* To reduce the variance of a spectral estimate *without any loss in frequency resolution, it is necessary to reduce **only the effect of the numerator*** and this can be achieved by dividing the signal transfer function by an estimate of the numerator. The GD domain achieves this operation without any singularity problem as it involves the multiplication of GD by the squared spectral magnitude of the numerator, rather than division.

If  $x(n)$  is a signal generated by an all-pole system driven by a white noise or sinusoids associated with white noise and further if it can be represented by  $X(\omega) = N(\omega)/D(\omega)$ , then  $D(\omega)$  corresponds to the system or sinusoids and  $N(\omega)$  to the

excitation or the associated noise. Its group delay is

$$\tau(\omega) = \tau_N(\omega) - \tau_D(\omega)$$

$$\tau(\omega) = \frac{K_N}{|N(\omega)|^2} - \frac{K_D}{|D(\omega)|^2} \quad (3a)$$

$\tau_N(\omega)$  and  $\tau_D(\omega)$  are group delay functions of  $N(\omega)$  and  $D(\omega)$ , respectively.  $K_N$  and  $K_D$  can be considered as constants for simplicity [3]. For the zeros close to unit circle due to signal truncation or input driving noise of a system or the associated noise with the signal,  $|N(\omega)|^2$  will be small. Hence, in the group delay domain, from Eq. (3a), the large magnitude ripples (due to first term) will mask the peaks (due to second term) as the magnitude of  $|D(\omega)|^2$  is sufficiently large for poles which are well inside the unit circle. On multiplying  $\tau(\omega)$  (Eq. (6)) [3,4] by  $|N(\omega)|^2$ ,

$$\tau_o(\omega) = \tau(\omega) |N(\omega)|^2$$

or

$$\tau_o(\omega) = K_N - \frac{K_D}{|D(\omega)|^2} |N(\omega)|^2 \quad (3b)$$

Here, in the first term, the effect of large values due to zeros close to the unit circle is removed and it will not mask the second term  $\tau_D(\omega)$ , which is due to poles. However,  $\tau_D(\omega)$  is multiplied by  $|N(\omega)|^2$ . But as the envelope of  $|N(\omega)|^2$  is nearly flat, the significant features of  $\tau_D(\omega)$  continue to exist, with the  $|N(\omega)|^2$  fluctuations superimposed on it. To get  $\tau_o(\omega)$ ,  $|N(\omega)|^2$  has to be determined with the given signal. The  $|N(\omega)|^2$  estimate,  $|\hat{N}(\omega)|^2$ , is derived using a cepstrally smooth spectrum  $|\tilde{X}(\omega)|^2$  obtained by truncating the cepstral coefficient sequence of the signal and is given by

$$|\hat{N}(\omega)|^2 = \frac{|X(\omega)|^2}{|\tilde{X}(\omega)|^2} \quad (3c)$$

### 2.2.3 2D group delay functions

As for the proposed DCT-based 2D-spectral estimator, the knowledge of 2D-group delay function is necessary, this section considers its details.

For a minimum phase signal (similar to Eqs. (2a) and (2b)),

$$\ln |X(\omega_1, \omega_2)| = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p(m, n) \cos(\omega_1 m + \omega_2 n) \quad (4a)$$

$$\theta(\omega_1, \omega_2) = \phi(\omega_1, \omega_2) + 2\pi I(\omega_1, \omega_2)$$

$$= - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p(m, n) \sin(\omega_1 m + \omega_2 n)$$

$$\theta(\omega_1, \omega_2) = - \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p(m, n) \sin(\omega_1 m + \omega_2 n) \quad (4b)$$

where  $\phi(\omega_1, \omega_2)$  is the principal value of phase and  $I(\omega_1, \omega_2)$  is an integer-valued function such that the overall phase

becomes a continuous function of  $\omega_1$  and  $\omega_2$ , and  $\theta(\omega_1, \omega_2)$  is the 2D-unwrapped minimum phase function. *Presently, the minimum phase  $\theta(\omega_1, \omega_2)$  is the phase derived from spectral magnitude  $|X(\omega_1, \omega_2)|$  using Eqs. (4a) and (4b) [6] and the minimum phase signal refers to that obtained from the spectral magnitude and the minimum phase.* The MGDs with respect to frequencies  $\omega_1$  and  $\omega_2$  are given by [6]

$$\tau_{m1}(\omega_1, \omega_2) = - \frac{\partial \theta(\omega_1, \omega_2)}{\partial \omega_1}$$

and

$$\tau_{m2}(\omega_1, \omega_2) = - \frac{\partial \theta(\omega_1, \omega_2)}{\partial \omega_2}$$

These two different MGDS with respect to frequencies  $\omega_1$  and  $\omega_2$  are given by

$$\tau_{m1}(\omega_1, \omega_2) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} m p(m, n) \cos(\omega_1 m + \omega_2 n) \quad (5a)$$

$$\tau_{m2}(\omega_1, \omega_2) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} n p(m, n) \cos(\omega_1 m + \omega_2 n) \quad (5b)$$

Given the two MGDs,  $\tau_{m1}(\omega_1, \omega_2)$  and  $\tau_{m2}(\omega_1, \omega_2)$  and if  $p_1(m, n)$   $p_2(m, n)$ ,  $m = 0, \dots, N-1$ ;  $n = 0, \dots, N-1$  are their corresponding two cepstral coefficient sequences, the combined cepstral sequence  $p(m, n)$  can be obtained. For this, the first row of  $p(m, n)$  consists of those derived from  $\tau_{m2}(\omega_1, \omega_2)$  and the first column of  $p(m, n)$  consists of those derived from  $\tau_{m1}(\omega_1, \omega_2)$ . The rest of the 2D array,  $p(m, n)$ , is computed as the average of  $p_1(m, n)$  and  $p_2(m, n)$  [6] as given below:

$$p(0, n) = p_2(0, n); \quad n = 1, \dots, N/2,$$

$$p(m, 0) = p_1(m, 0); \quad m = 1, \dots, N/2$$

$$p(m, n) = \frac{1}{2} (p_1(m, n) + p_2(m, n)),$$

$$m = 1, \dots, N/2; \quad n = 1, \dots, N/2$$

$$p(m, n) = 0, \quad m = N/2 + 1 : N-1;$$

$$n = N/2 + 1 : N-1 \quad (5c)$$

The 2D log spectrum is obtained by taking the 2DFT of  $p(m, n)$ , i.e.

$$\ln |X(\omega_1, \omega_2)| = \text{DFT}[p(m, n)] \quad (5d)$$

### 3 The proposed 2D-spectral estimator based on 2D-analytic DCT and 2D-modified magnitude group delay (2D-MMGD)

In 1D-spectral estimation [4], the MMGD was applied to the analytic DCT derived from the DCT of the original signal. Hence, for the proposed 2D-spectral estimation, it is necessary to apply the 2D-MMGD to the 2D-analytic DCT. The 2D-analytic DCT not only involves 2D-DCT of the signal

**Fig. 2** **a**  $X(k_1, k_2)$ , **b**  $X_a(k_1, k_2)$ , **c** method of computing  $X_a(k_1, k_2)$ , **d**  $H_2(k_1, k_2)$ , **e**  $H_3(k_1, k_2)$ , **f**  $H_4(k_1, k_2)$

but also 2D-Hilbert transform of the 2D-DCT. Hence, the 2D-Hilbert transform and the 2D-MMGD are required in realizing the proposed 2D-spectral estimation and will be considered.

The Hilbert transform of a signal can be achieved by setting the magnitude of the negative frequency components of the DFT to zero and taking the imaginary part of its inverse DFT. However, such a Hilbert transform will suffer from DFT leakage, as during the computation of the DFT, there will be leakage of energy from negative frequency components to positive frequency components and vice-versa. This can be solved by convolving the signal with the impulse response of the Hilbert transformer, in time domain. The transfer function and the impulse response of the Hilbert transformer are

$$H(\omega) = \begin{cases} -j & \omega > 0 \\ j & \omega \leq 0 \end{cases} \quad h(n) = \frac{2 \sin^2(\pi n/2)}{n\pi}, \quad -\infty \leq n \leq \infty$$

Thus, to generate an accurate analytic signal, a Hilbert transformer impulse response is essential.

For a causal 2D-signal also, making the 2D-DFT values as zeros in all the three quadrants, except in its first quadrant, and taking its inverse 2D-DFT can result in its analytic signal. However, similar to 1D-signal case, there will be leakage from all the three quadrants to the first quadrant and vice versa, once the 2D-DFT is taken, and hence the analytic 2D-signal derived by this process will not be accurate. To get an accurate 2D-analytic signal, the imaginary part of the analytic signal, which is the Hilbert transform of the real part, has to be derived by convolving the signal with the impulse response of the Hilbert transformer.

### 3.1 2D-Hilbert transformer

The development of the 2D-Hilbert transformer is based on the concept that for an analytic signal  $x_a(n_1, n_2)$ , its Fourier

transform  $X_a(k_1, k_2)$  will be confined to only the first quadrant and it will be zero in all other three quadrants. This can be represented as

$$x_a(k_1, k_2) = 0, \quad k_1 < 0, k_2 < 0 \quad (6a)$$

where  $k_1$  and  $k_2$  are the spatial frequencies involved for a 2D-signal. Further, the imaginary part of the Fourier transform forms the Hilbert transform of the analytic signal. This approach is similar to the method considered in [7]. Since the Fourier transform is confined to only the first quadrant, it is not conjugate symmetric, which in turn says that the analytic sequence is complex. Hence, the objective is to find an operation which when applied on  $x(n_1, n_2)$  would yield  $x_i(n_1, n_2)$  which when multiplied by  $j$  and added to  $x(n_1, n_2)$  would give an analytic sequence,

$$x_a(n_1, n_2) = x(n_1, n_2) + jx_i(n_1, n_2) \quad (6b)$$

For convenience, it is desirable to visualize the DFT of  $x(n_1, n_2)$ ,  $X(k_1, k_2)$  as given in Fig. 2a. Here, the first quadrant contents of the sequence are denoted by A, second quadrant by B and so on. Now, given  $X(k_1, k_2)$ , it is required to get a sequence  $X_a(k_1, k_2)$  (Fig. 2b), and a method of achieving this is shown in Fig. 2c.

The second, third and fourth components in the left-hand side of the equation in Fig. 2c can be obtained by multiplying  $X(k_1, k_2)$  by the sequences  $H_2(k_1, k_2)$ ,  $H_3(k_1, k_2)$  and  $H_4(k_1, k_2)$  shown in Fig. 2d–f, respectively. Note that  $H_2(k_1, k_2)$ ,  $H_3(k_1, k_2)$  and  $H_4(k_1, k_2)$  are separable sequences; i.e.,

$$H_2(k_1, k_2) = \text{sgn}(k_1)un(k_2)$$

$$H_3(k_1, k_2) = un(k_1)\text{sgn}(k_2)$$

$$H_4(k_1, k_2) = \text{sgn}(k_1)\text{sgn}(k_2)$$

$$\text{where, } un(k) = 1, \forall k, \text{ and } \text{sgn}(k) = \begin{cases} +1, & k \geq 0 \\ -1, & k < 0 \end{cases}$$



The IDFT of  $un(k)$  and  $\text{sgn}(k)$ , respectively, is,

$$\text{IDFT}[un(k)] = \delta(n),$$

$$\text{IDFT}[\text{sgn}(k)] = \begin{cases} j \frac{2}{\pi} \frac{\sin^2(\pi n/2)}{n} & n \neq 0 \\ 0 & n = 0 \end{cases}$$

By applying the property of separable sequences,

$$\begin{aligned} h_2(n_1, n_2) &= \text{IDFT}[H_2(k_1, k_2)] \\ &= \text{IDFT}[\text{sgn}(k_1)] \text{IDFT}[un(k_2)] \end{aligned} \quad (7a)$$

$$\begin{aligned} h_3(n_1, n_2) &= \text{IDFT}[H_3(k_1, k_2)] \\ &= \text{IDFT}[un(k_1)] \text{IDFT}[\text{sgn}(k_2)] \end{aligned} \quad (7b)$$

$$\begin{aligned} h_4(n_1, n_2) &= \text{IDFT}[H_4(k_1, k_2)] \\ &= \text{IDFT}[\text{sgn}(k_1)] \text{IDFT}[\text{sgn}(k_2)] \end{aligned} \quad (7c)$$

Here,  $h_2(n_1, n_2)$  and  $h_3(n_1, n_2)$  are imaginary, and  $h_4(n_1, n_2)$  is real. Now,

$$\begin{aligned} jX_i(k_1, k_2) \\ &= X(k_1, k_2)[H_2(k_1, k_2) + H_3(k_1, k_2) + H_4(k_1, k_2)] \end{aligned}$$

So,

$$\begin{aligned} jx_i(n_1, n_2) \\ &= x(n_1, n_2)^*[h_2(n_1, n_2) + h_3(n_1, n_2) + h_4(n_1, n_2)] \\ &= x(n_1, n_2)^*h(n_1, n_2) \end{aligned} \quad (8)$$

where  $*$ : 2D convolution and

$$h(n_1, n_2) = [h_2(n_1, n_2) + h_3(n_1, n_2) + h_4(n_1, n_2)] \quad (9)$$

Then,

$$x_a(n_1, n_2) = x(n_1, n_2) + jx_i(n_1, n_2) \quad (10)$$

Here,  $h(n_1, n_2)$  (Eq. 9) is the desired 2D Hilbert transform impulse response and  $x_a(n_1, n_2)$  (Eq. 10) is the analytic signal corresponding to  $x(n_1, n_2)$ .

### 3.2 2D-analytic DCT from 2D-DCT using 2D-Hilbert transform

In 1D signal spectral estimation using the DCT [4], the analytic DCT,  $C_a(k)$  is computed and this has the desired properties of DCT as its real and imaginary parts are  $C(k)$ , DCT of the signal and  $\hat{C}(k)$ , Hilbert transform of  $C(k)$ , respectively. That is,

$$C_a(k) = C(k) + j\hat{C}(k)$$

The analytic DCT is nothing but the DFT of the one-sided signal. This is based on the fact that for an analytic signal, the spectrum is one-sided and for the one-sided signal, the spectrum is analytic. Thus, the analytic DCT forms the spectrum or DFT of one-sided signal. But this DFT is having the desired properties of DCT as its components are derived from DCT [4], as already mentioned.

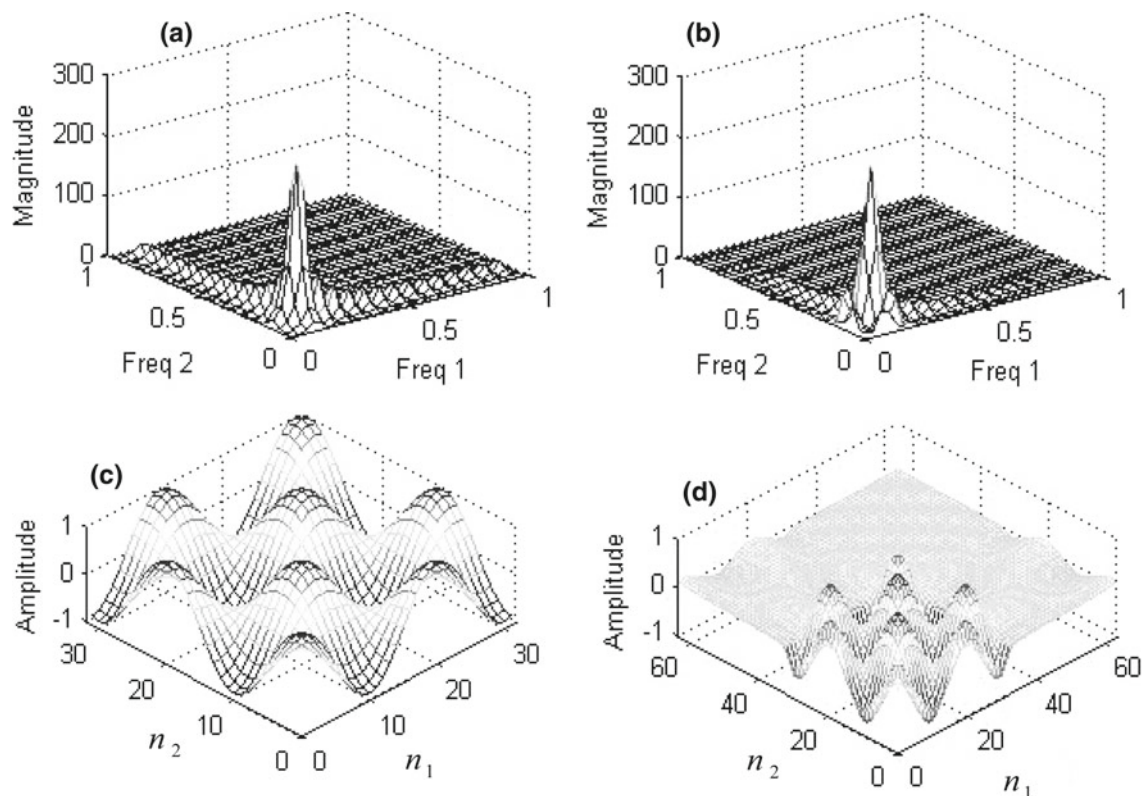
The analytic DCT of  $x(n_1, n_2)$  can be obtained from its DCT  $C_x(k_1, k_2)$  by computing its analytic DCT using a 2D-Hilbert transform, explained already. Such a 2D-analytic DCT forms the transform of a 2D signal  $y(n_1, n_2)$ , which has nonzero values only in the first quadrant and zero values in all other quadrants. This 2D-analytic signal is of dimension  $2N_1 \times 2N_2$  and has a linear phase shift of  $1/2$  the sampling interval in both axes (and this is due to symmetry with even number of points with the absence of central point for DCT-II). This does not pose any problem as the interest is only 2D power spectral estimation which considers only the spectral magnitude and not the phase spectrum of the signal.

As discussed before, the  $N_1 \times N_2$ -point DCT can be interpreted as the  $2N_1 \times 2N_2$ -point DFT of the appropriately extended signal with a linear phase corresponding to half the sampling period added to both the axes, and then only the first quadrant is selected. This  $N_1 \times N_2$ -point DCT is symmetrically extended to second, third and fourth quadrants to construct  $2N_1 \times 2N_2$ -point DCT using the symmetry property of the DCT [1]. The resulting  $2N_1 \times 2N_2$ -point DCT is convolved with 2D Hilbert transformer impulse response  $h(n_1, n_2)$  (Eq. 9) to get the equivalent  $2N_1 \times 2N_2$ -point Hilbert transform of the actual signal. The  $2N_1 \times 2N_2$ -point DCT and its Hilbert transform form the real and imaginary parts of the desired analytic signal. This will be reflected in the space domain (inverse 2D-DFT) as the reconstruction of  $x(n_1, n_2)$  from its extended version  $y(n_1, n_2)$ .

This method is illustrated using a signal of size  $32 \times 32$ ,  $x(n_1, n_2) = \cos(0.1\pi n_1) \cos(0.1\pi n_2)$ . The 2D-DCT of this signal is taken and then the 2D-analytic DCT is computed using the 2D Hilbert transform. The magnitude of the 2D-DFT of the actual signal and that of the 2D-analytic DCT are, respectively, shown in Fig. 3a, b. As expected, the unwanted high-frequency components along the two axes are significantly reduced for the 2D-analytic DCT compared to that of 2D-DFT. The actual signal is shown in Fig. 3c, and the IDFT of the 2D-analytic DCT is shown in Fig. 3d. Note that the resultant signal is essentially  $64 \times 64$ , which is the zero padded version of the actual signal.

### 3.3 2D-spectral estimation based on 2D-analytic DCT and MMGD

The 2D analytic DCT derived from the 2D-DCT can be used for spectral estimation. Similar to 1D modified magnitude group delay [2,3], the definition of a modified magnitude group delay (MMGD) in 2D facilitates the application of the MMGD to 2D-spectral estimation. This makes the 2D-spectral estimation immune to the driving noise of a 2D system/ associated noise with the 2D signal, as the MMGD removes the zeros close to the unit sphere which



**Fig. 3** **a** DFT of actual signal, **b** reconstructed DFT from the DCT, **c** actual signal, **d** IDFT of the reconstructed DFT

are due to them. This enables improvement in variance over that obtained by the 2D-DCT preserving its frequency resolution.

### 3.3.1 Extension of modified MGD for a 2D-signal

The extension of 1D-MMGD to 2D has been made for the purpose of 2D-spectral estimation and this will be considered.

If  $y(m, n)$  is a signal generated by an all-pole system, driven by a white noise and its spectrum is  $Y(\omega_1, \omega_2) = N(\omega_1, \omega_2)/D(\omega_1, \omega_2)$ , then  $D(\omega_1, \omega_2)$  corresponds to the system or sinusoids and  $N(\omega_1, \omega_2)$  to the excitation or the associated noise. The MGD with respect to  $\omega_1$  is

$$\tau_{m1}(\omega_1, \omega_2) = \tau_{m1N}(\omega_1, \omega_2) - \tau_{m1D}(\omega_1, \omega_2) \quad (11a)$$

$\tau_{m1N}(\omega_1, \omega_2)$  and  $\tau_{m1D}(\omega_1, \omega_2)$  are the MGDs for  $N(\omega_1, \omega_2)$  and  $D(\omega_1, \omega_2)$ , respectively, computed with respect to  $\omega_1$ . Also,  $\tau_{m1}(\omega_1, \omega_2)$  can be written as,

$$\tau_{m1}(\omega_1, \omega_2) = \frac{\alpha_N}{|N(\omega_1, \omega_2)|^2} - \frac{\alpha_D}{|D(\omega_1, \omega_2)|^2} \quad (11b)$$

Here,  $\alpha_N$  and  $\alpha_D$  are constants. Similar to 1D case, to remove the effect of zeros close to the unit sphere on  $\tau_{m1D}(\omega_1, \omega_2)$ , it is required to multiply  $\tau_{m1}(\omega_1, \omega_2)$  by

$|N(\omega_1, \omega_2)|^2$ . Hence, the modified MGD (MMGD) with respect to  $\omega_1$ ,  $\tau_{m1o}(\omega_1, \omega_2)$  is,

$$\begin{aligned} \tau_{m1o}(\omega_1, \omega_2) &= \tau_{m1}(\omega_1, \omega_2) |N(\omega_1, \omega_2)|^2 \\ \tau_{m1o}(\omega_1, \omega_2) &= \alpha_N - \frac{\alpha_D}{|D(\omega_1, \omega_2)|^2} |N(\omega_1, \omega_2)|^2 \end{aligned} \quad (12a)$$

Here, similar to 1D case, the effect of large values due to zeros close to the unit sphere is removed and it will not mask the second term  $\tau_{m1D}(\omega_1, \omega_2)$ , which is due to poles. The estimate of  $|N(\omega_1, \omega_2)|^2$  is given by

$$|\tilde{N}(\omega_1, \omega_2)|^2 = |Y(\omega_1, \omega_2)|^2 / |\bar{Y}(\omega_1, \omega_2)|^2 \quad (12b)$$

$|\bar{Y}(\omega_1, \omega_2)|^2$  is the cepstrally smoothed power spectrum obtained by only the initial few coefficients of the cepstral sequence. Similarly,

$$\tau_{m2o}(\omega_1, \omega_2) = \tau_{m2}(\omega_1, \omega_2) |N(\omega_1, \omega_2)|^2 \quad (13a)$$

Two different sets of causal cepstral coefficient sequences,  $p_1(m, n)$  and  $p_2(m, n)$  obtained from the MMGDs  $\tau_{m1o}(\omega_1, \omega_2)$  and  $\tau_{m2o}(\omega_1, \omega_2)$ , respectively, are combined to get  $p(m, n)$  by the procedure already explained (Eqn.(5c)). From these cepstral coefficients  $c(m, n)$ , the magnitude spectrum can be reconstructed.

$$\ln |Y(\omega_1, \omega_2)| = DFT[p(m, n)] \quad (13b)$$

This 2D-spectrum will be significantly less affected by the driving noise to a 2D-system whose output is the signal or noise associated with the 2D-signal or truncation effects and hence will have less variance but preserves the frequency resolution of the rectangular window.

If the MGDs  $\tau_{m1}(\omega_1, \omega_2)$  and  $\tau_{m2}(\omega_1, \omega_2)$  are computed using the 2D-DFT, the spectral estimation method is called DFT-MMGD method. If the MGDs  $\tau_{m1}(\omega_1, \omega_2)$  and  $\tau_{m2}(\omega_1, \omega_2)$  are computed from the 2D-analytic DCT using the 2D Hilbert transform, the spectrum estimation method is called DCT-MMGD method.

#### 4 Simulation results

The proposed two methods of 2D-spectral estimation methods, DFT-MMGD and DCT-MMGD, each are applied to different simulated signals viz., sinusoids plus white Gaussian noise, AR process plus white Gaussian noise and two low-level 2D sinusoids in the presence of an 2D-AR process having high-level spectral peaks. The results are also compared with those of 2D-LP method [5]. The mean and the variance of the 2D-spectrum are used as indices of performance. The mean spectrum  $S_m(k_1, k_2)$  is given by

$$S_m(k_1, k_2) = \frac{1}{P} \sum_{i=1}^P S_i(k_1, k_2) \quad (14a)$$

$S_i(k_1, k_2)$  is the individual spectrum estimate for the  $i^{th}$  trial and  $P$  is the number of trials ( $P$  different signals) considered in the ensemble. The mean spectrum is supposed to be close to or it should bring out all the desired features of the true spectrum like resolving spectral peaks. The variance of the method of spectrum estimate indicates the variability in the spectrum estimate from its mean spectrum for different trial segments of signals. The variance of the spectrum,  $\text{var}(k_1, k_2)$ , is given by

$$\begin{aligned} \text{var}(k_1, k_2) &= \frac{1}{P} \sum_{i=1}^P [S_i(k_1, k_2) - S_m(k_1, k_2)]^2 \\ 0 &\leq k_1 \leq N_1 - 1 \\ 0 &\leq k_2 \leq N_2 - 1 \end{aligned} \quad (14b)$$

The variance spectrum being a surface for a 2D signal, it is difficult to quantify the variance performance at all frequencies (in both dimensions) by looking at the surface. This necessitates for a single performance index, which accounts for the variance at all frequencies simultaneously. Hence, the average variance per frequency bin (AVPFB)  $\rho$  which quantifies the variance by a single parameter has been used and

is of importance which is given by

$$\rho = \frac{1}{N_1 N_2} \sum_{k_1=N_2}^{N_1} \sum_{k_2=N_1}^{N_2} \text{var}(k_1, k_2) \quad (14c)$$

where  $(N_1 \times N_2)$  is spectrum size. In situations like sinusoids plus noise, the true spectrum is not known unlike in the case of AR process where by knowing the actual coefficients used for generating the AR process, it is possible to get the true spectrum, and the root mean square error (RMSE)/mean square error (MSE) can be used as a performance index. The MSE is related to variance and bias (the difference between the true spectrum and the mean spectrum). The mean square error is equal to bias square plus variance. In many cases, it is of interest that the mean spectrum brings out the important characteristics of the true spectrum even though the former does not have perfect match to the latter (though such a match is highly desirable) and the variance of the spectrum estimate is less. This evaluation approach is generally followed, and in the present study, the performance is evaluated from this point of view (which is practical for sinusoids plus noise example as its true spectrum is impossible to be known). Especially with real data, the true spectrum will not be known and only the mean spectrum and variance spectrum decide the quality of the spectral estimate.

The 2D data size considered is  $(128 \times 128)$  and is sufficient to resolve the spectral peaks considered in this study. The data size is an important specification for spectral estimation using DFT/DCT as it decides frequency resolution. Further, for estimating the variance, it is desirable to consider number of trials  $P$ , very large. Presently,  $P = 25$  is used, and comparison of the variance of different methods is made for the same value of  $P$ .

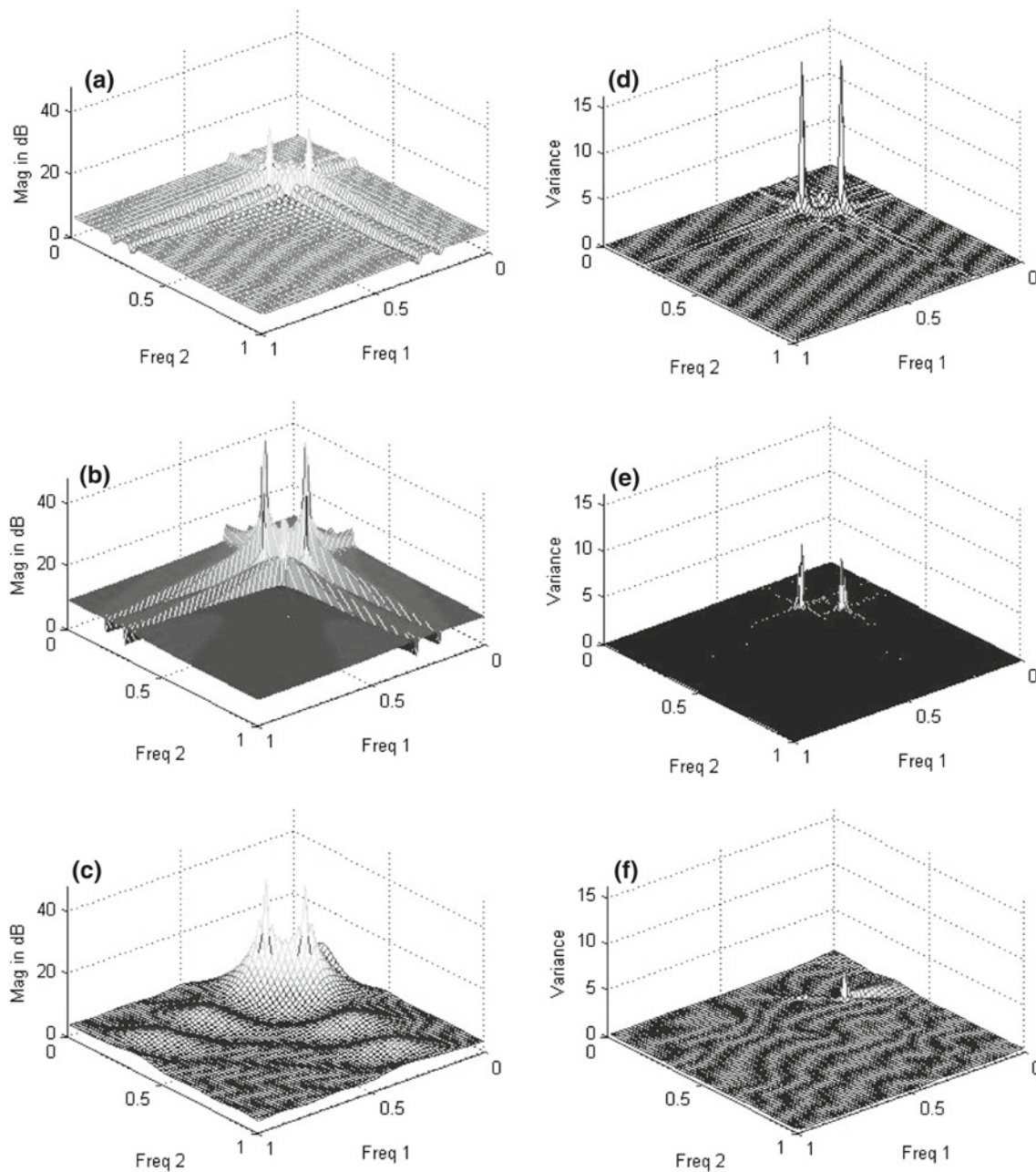
(a) *Spectral estimation for sinusoids plus noise input:* The 2D-sinusoids considered with noise are given by

$$\begin{aligned} x(n_1, n_2) &= \cos(0.2\pi n_1 + \theta(i)) \cos(0.3\pi n_2 + \theta(i)) \\ &\quad + \cos(0.3\pi n_1 + \theta(i)) \cos(0.2\pi n_2 + \theta(i)) \\ &\quad + w(n_1, n_2) \end{aligned}$$

where  $w(n_1, n_2)$ , the zeros mean white Gaussian noise with unit variance, and  $\theta(i)$ , the phase that changes randomly from trial to trial.

The 2D-cepstrum used is of size  $(4 \times 4)$ . For the 2D-LP method, the order used is 8. For a  $SNR = 20\text{dB}$ , the mean spectra of DFT-MMGD, DCT-MMGD and 2D-LP methods are shown in Fig. 4a–c, respectively. Figure 4d–f show the variances of 4a–c, respectively. It is seen that from the mean spectra, the DCT-MMGD resolves the two spectral peaks very well (height of the spectral peaks above the common base level) compared to those of other two



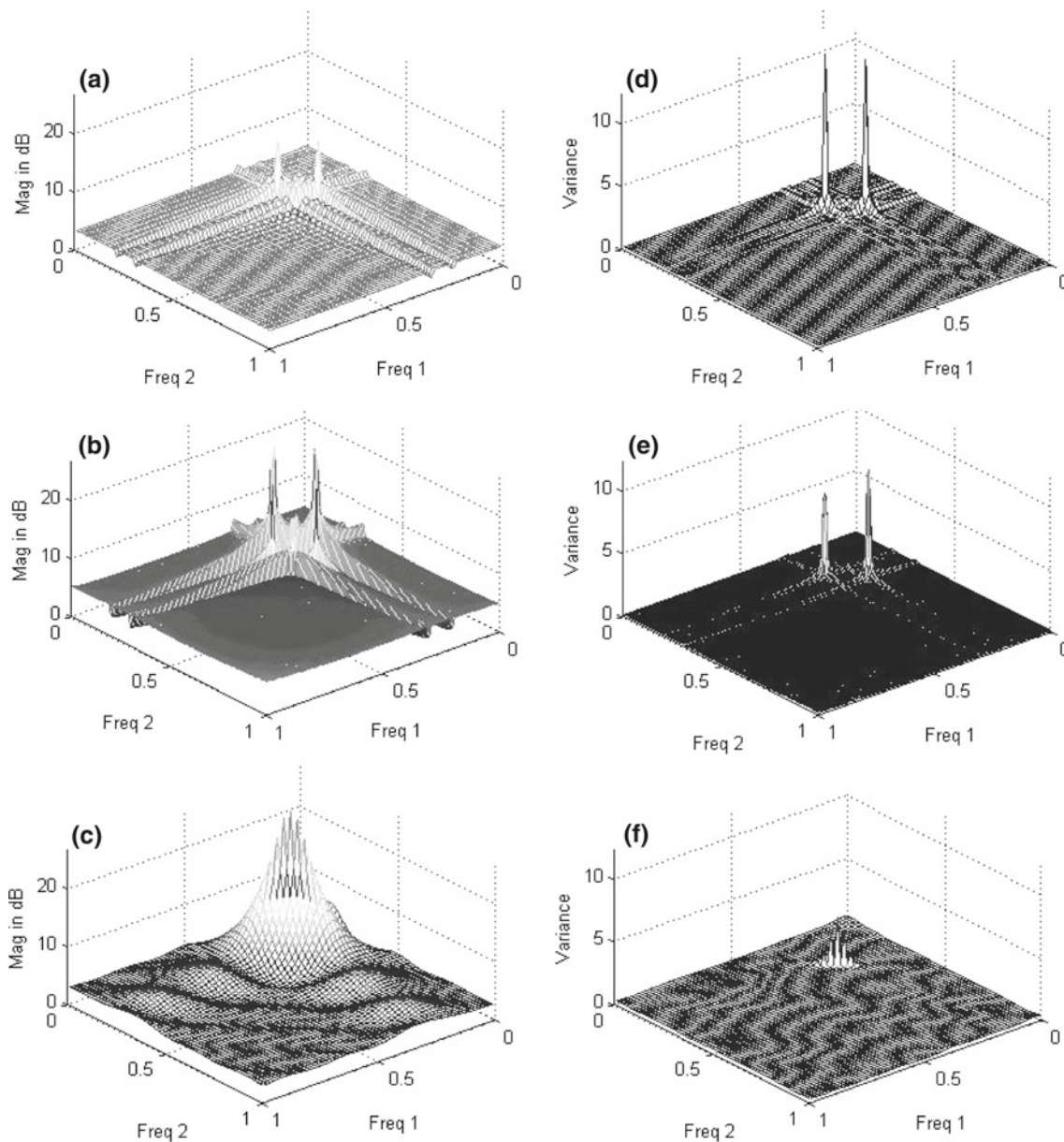


**Fig. 4** Mean and variance spectra for 2D-sinusoids at an SNR=20 dB. Mean spectra: **a** DFT-MMGD, **b** DCT-MMGD and **c** 2D-LP. Variance spectra: **d** DFT-MMGD, **e** DCT-MMGD and **f** 2D-LP

methods. Further, in this respect, the 2D-LP method is better than the DFT-MMGD. From the variance spectra, it is clear that DCT-MMGD has a significantly better variance than that of the DFT-MMGD (by a factor more than 2). The variance achieved by 2D-LP method is significantly superior to those of the two methods. But the overall performance of the DCT-MMGD is better than 2D-LP method as the former's resolution is good and the variance is tolerable.

The performance of the methods at SNR = -5 dB is shown in Fig. 5. It is seen that from the mean spectra

(Fig. 5a, b), for the DFT-MMGD and DCT-MMGD, the spectral peaks are well resolved but the heights of the spectral peaks are reduced approximately by a factor of half, indicating that the frequency resolution is reduced due to the increased noise level and this is expected. At this poor SNR, the 2D-LP method fails to resolve the two peaks as it shows only a single broad peak covering both the peaks (Fig. 5c). At the spectral peak locations, the variance of the DFT-MMGD and DCT-MMGD (Fig. 5d, e) are of the same order as at SNR=20dB (Fig. 4d, e), indicating that they perform better at low SNR, in terms of variance. As the 2D-LP method is



**Fig. 5** Mean and variance spectra for 2D sinusoidal estimates for SNR = -5 dB. Mean spectra: **a** DFT-MMGD, **b** DCT-MMGD and **c** 2D-LPB. Variance spectra: **d** DFT-MMGD, **e** DCT-MMGD and **f** 2D-LPB

unable to resolve the two spectral peaks (Fig. 5c, f), its lower variance is of no value.

The AVPFB for the three methods is listed for SNR ranging from -5 to 20 dB in Table 1. The AVPFB for the DCT-MMGD is significantly better than that of DFT-MMGD. Further, its value is comparable to that of the 2D-LP method in the SNR range of 0–20 dB. But for both MMGD methods, the AVPFB increases as the SNR decreases which is expected, and this is against the local variance stated as the AVPFB provides a global feature covering the complete 2D-frequency range. The above results indicate that the 2D-LP method cannot perform well in the presence of noise. This is due to the

**Table 1** AVPFB for sinusoids

SNR (dB)	DFT-MMGD	DCT-MMGD	2D-LP
20	0.0538	0.0147	0.149
10	0.0205	0.0118	0.131
5	0.0304	0.0134	0.137
0	0.0361	0.0165	0.127
-5	0.0398	0.0314	0.124

fact that the 2D-LP method is basically dependent on the signal correlation property. However, the presence of noise will disturb the desired correlation property for the

LP method and hence as expected its performance is poor.

(b) *Spectral estimation for the autoregressive process plus noise input:*

For a 2D-AR process, a 2D denominator  $D(z_1, z_2)$  is generated from two 1D polynomials  $A_1(z_1)$  and  $A_2(z_2)$  as

$$D(z_1, z_2) = A_1(z_1)A_2(z_2)$$

$$A_1(z) = 1 - 3.0672z^{-1} + 4.2441z^{-2} - 3.0062z^{-3} + 0.9606z^{-4}$$

and

$$A_2(z) = 1 - 1.5168z^{-1} + 0.9801z^{-2}$$

The roots of  $A_1(z_1)$  are  $0.99e^{\pm 25\pi/180}$ ,  $0.99e^{\pm 50\pi/180}$  and of  $A_2(z_1)$  are  $0.99e^{\pm 40\pi/180}$ .

The 2D-AR process is generated by filtering a 2D-Gaussian white noise of zero mean and unit variance by the transfer function

$$C(z_1, z_2) = 1/D(z_1, z_2)$$

In this case also the DFT-MMGD, DCT-MMGD and 2D-LP methods are applied. In both DFT-MMGD and the DCT-MMGD, the cepstral coefficient truncation length used is  $(4 \times 4)$ . The performance of DCT-MMGD in terms of spectral resolution is almost similar to that of DFT-MMGD (Fig. 6a, b). But in terms of variance, the DCT-MMGD is significantly better than DFT-MMGD (Fig. 6d, e). Compared to these methods, the 2D-LP mean spectrum brings out the spectral peaks very well and is having close resemblance to the true AR spectrum obtained by the 2D-AR coefficient polynomial. Further, its variance is also very significantly smaller than those by the other two methods. The order for the 2D-LP method used is 12.

For the same AR process with a SNR=0 dB, the results are shown in Fig. 7. The mean spectrum by the DCT-MMGD method brings out the two spectral peaks better than that by the DFT-MMGD (Fig. 7a, b). But the variance at the spectral peak locations by the two methods is comparable (Fig. 7d, e). The 2D-LP resolves the spectral peaks, but the peaks are significantly broader than those achieved by the two new methods indicating their superior resolving power. Further, its variance is considerable (Fig. 7f) compared to that achieved for sinusoids even with SNR = -5 dB (Fig. 5f).

The AVPFB for different SNR values ranging from 20 to -5 dB is given in Table 2. At high SNR, the AVPFB values achieved by the DCT-MMGD method are less than those of DFT-MMGD method. Further, as the SNR decreases, AVPFB values decrease indicating they have better global variance performance at low SNR than at high SNR. For the 2D-LP method, the AVPFB values are almost same irrespective of SNR and are more than those achieved by the

new methods indicating their improved global variance over that of the 2D-LP method. At SNR = -5 dB, the AVPFB of the DFT-MMGD becomes less than that of the DCT-MMGD method.

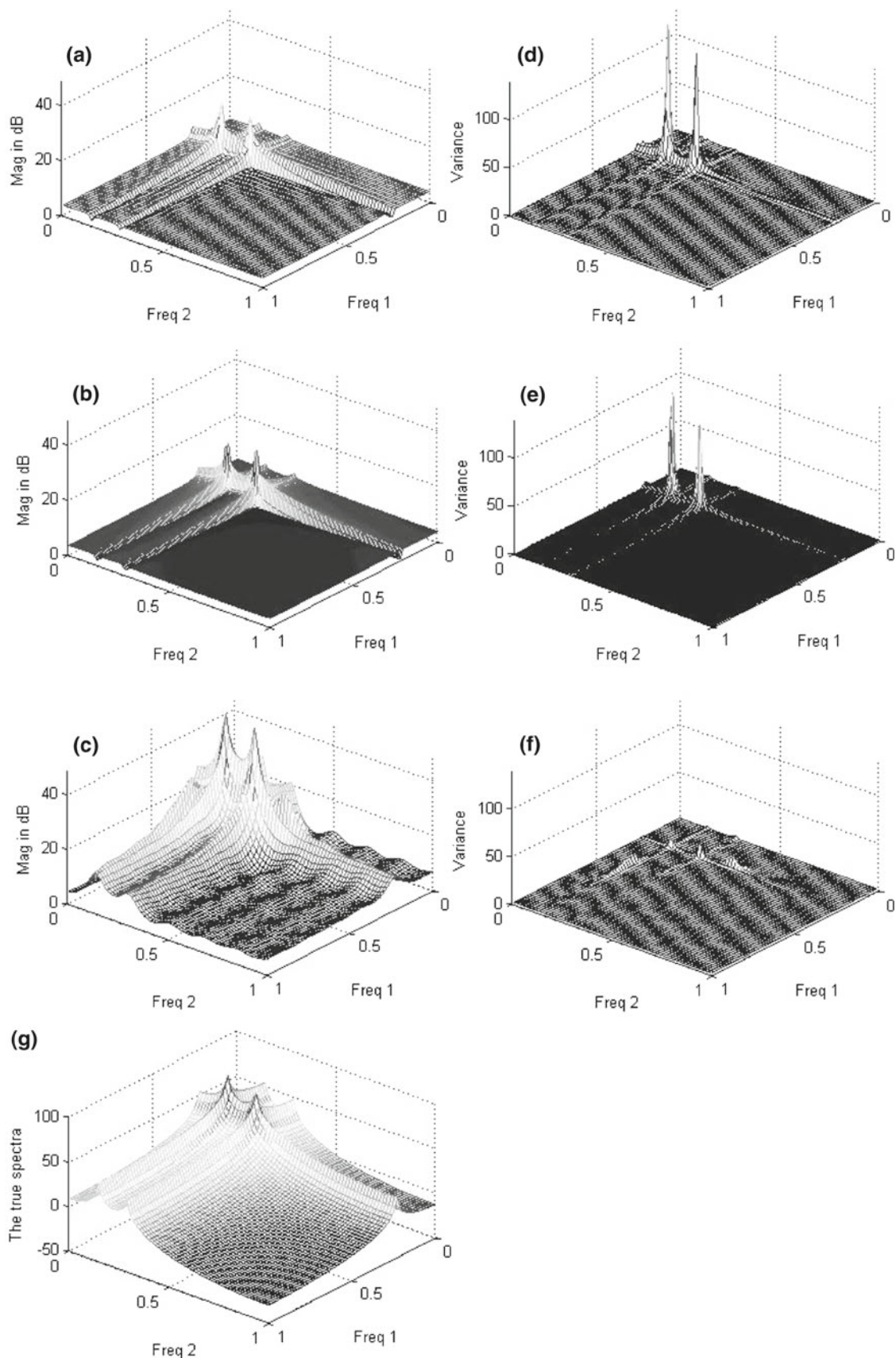
(c) *Detection performance of the proposed Spectral estimation methods for low-level sinusoids in the presence of a 2D-AR process with high-level peaks:*

In the 1D-DFT spectrum, the leakage from large amplitude peaks may mask the low-amplitude peaks. But as this effect is significantly less in the DCT, DCT-MMGD will have a better detectability of the low-amplitude peaks compared to that of DFT-MMGD method [4, 8]. To illustrate this to the 2D-AR process (already considered), 2D sinusoids having peaks at frequencies  $(0.6\pi, 0.65\pi)$  and  $(0.65\pi, 0.6\pi)$  are added and further to this, a 2D Gaussian white noise is added to have an SNR = 25 dB. The amplitude of sinusoids is such that their spectral peak magnitudes are 20 dB below the spectral peaks of the AR process. The mean spectra obtained using 25 realizations for the DFT-MMGD and DCT-MMGD are shown in the Fig. 8a, b. The low-amplitude peaks are better brought out by DCT-MMGD compared to the DFT-MMGD method. However, for the 2D-LP method, from the mean spectrum, the presence of these two low-level spectral peaks is not explicitly evident. The contour plots for the low-level spectral peak region indicate that the two peaks are detected by both DFT-MMGD and DCT-MMGD methods (Fig. 8d, e). Further, the DCT-MMGD detects them with a better frequency resolution than by the DFT-MMGD. Though for the 2D-LP method, the contour plot indicates the presence of these two low-level peaks (which is not clear from the magnitude spectrum (Fig. 8c) but with a poorer frequency resolution as the inner contour for one of the peaks is broader. But the new methods are able to detect with better frequency resolution, and in this, the DCT-MMGD is superior due to DCT inherent higher-frequency resolution compared to that of DFT. In both DFT-MMGD and the DCT-MMGD, the cepstral coefficient truncation length used is  $(4 \times 4)$  and the order of 2D-LP used is 12.

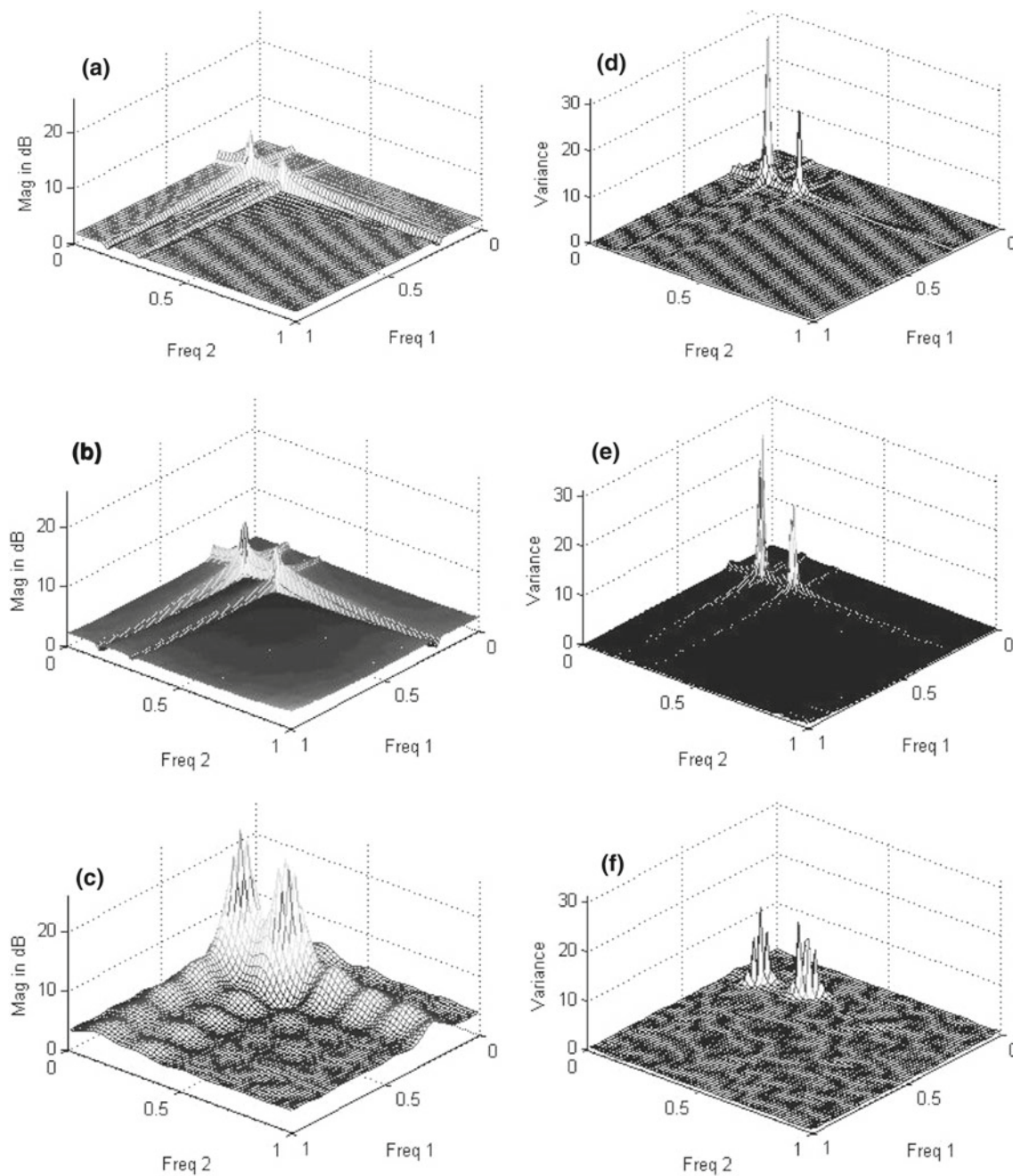
## 5 Conclusions

In this paper, two new methods of 2D-spectral estimation methods were presented. In the first method, the 2D-modified magnitude group delay (MMGD), which is an extension from 1D to 2D, is applied to the 2D-DFT (DFT-MMGD). In the second, the MMGD is applied to 2D-analytic DCT (DCT-MMGD) which is realized by a 2D-DCT and its Hilbert transform. Due to the DCT desirable properties like improved frequency resolution and low leakage compared to those of DFT and due to the MMGD ability of reducing variance preserving the original frequency resolution, the DCT-MMGD method provided a better spectral estimate





**Fig. 6** Mean and variance spectra for 2D AR process at SNR=20 dB. Mean spectra: **a** DFT-MMGD, **b** DCT-MMGD, and **c** 2D-LPB. Variance spectra: **d** DFT-MMGD, **e** DCT-MMGD, **f** 2D-LP, and **g** true spectrum



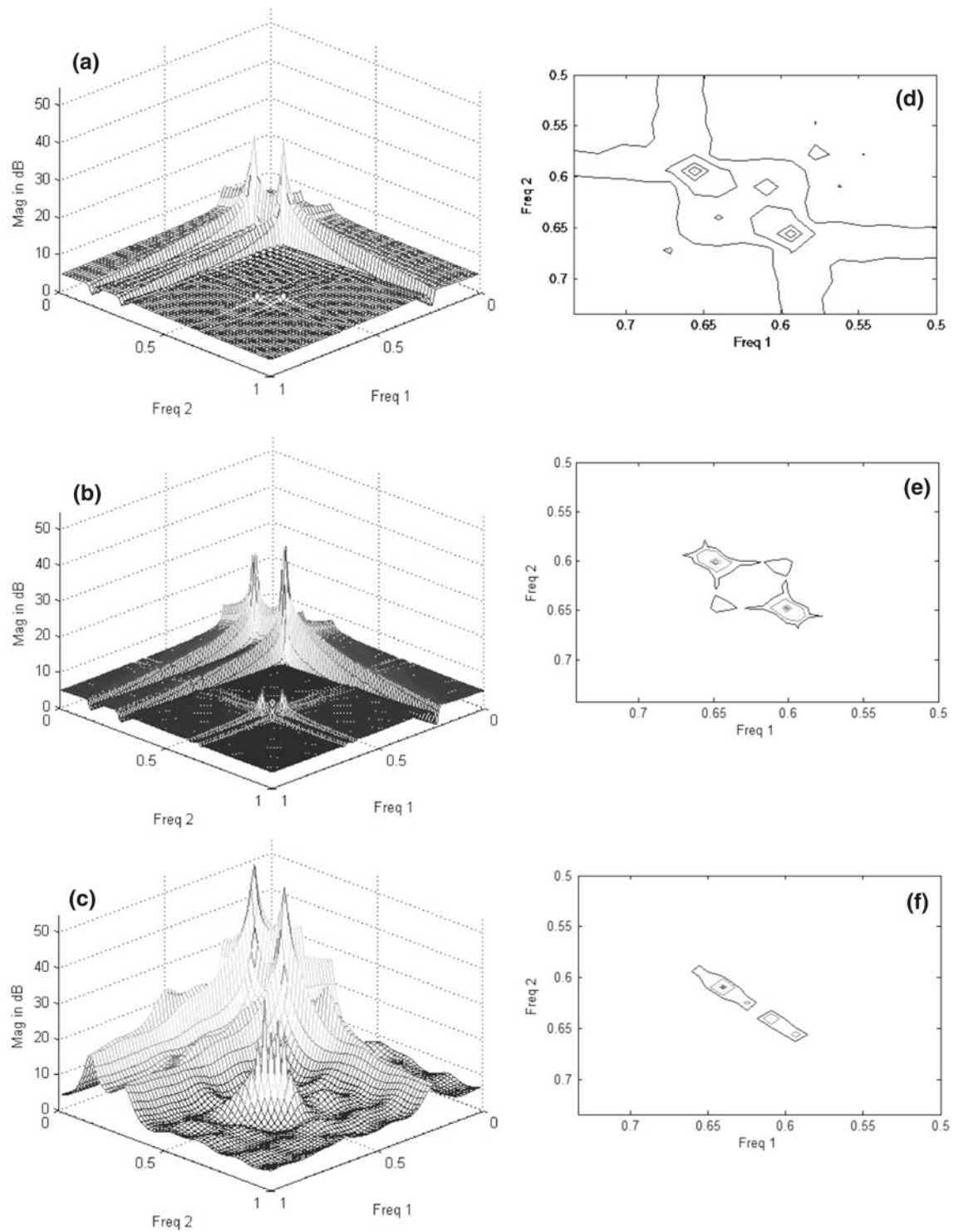
**Fig. 7** Mean and variance spectra for 2D AR process at SNR=0 dB. Mean spectra: **a** DFT-MMGD, **b** DCT-MMGD, and **c** 2D-LP. Variance spectra: **d** DFT-MMGD, **e** DCT-MMGD and **f** 2D-LP

which has better frequency resolution, variance and detectability of low-level spectral peaks, than by the DFT-MMGD. Further, as the MMGD removes the zeros close to the unit sphere due to noise, the two new methods performed better than the 2D-linear prediction method, in terms of frequency resolution, variance (both in local and global) and detectability. In this respect again, the DCT-MMGD outperforms DFT-MMGD. Due to MMGD nature, the new methods perform better at low SNR than at high SNR.

**Table 2** AVPFB for AR process

SNR (dB)	DFT-MMGD	DCT-MMGD	2D-LP
20	0.63	0.35	0.36
10	0.28	0.19	0.34
5	0.17	0.14	0.32
0	0.10	0.10	0.31
-5	0.07	0.18	0.32





**Fig. 8** Detection of low-level sinusoids in the presence of a high level of 2D AR process associate with noise SNR=25dB. Mean spectra: **a** DFT-MMGD, **b** DCT-MMGD, and **c** 2D-LP contour plots: **d–f** for the low-amplitude 2D sinusoids region of (a)–(c)

## References

1. Lim, J.S.: Two-dimensional Signal and Image Processing. Prentice Hall, Englewood Cliffs (1990)
2. Murthy, H.A., Yegnanarayana, B.: Speech processing using group delay functions. *Signal Process.* **22**, 259–267 (1991)
3. Yegnanarayana, B., Murthy, H.A.: Significance of group delay functions in spectral estimation. In: *IEEE Trans. Signal Process.* **40**, 2281–2289 (1992)
4. Narasimhan, S.V., Harish, M.: Spectral estimation based on discrete cosine transform and modified group delay. *Signal Process.* **86**, 1586–1596 (2006)
5. Lawrence Marple, S. Jr.: Two-dimensional lattice linear predictor parameter estimation method and fast algorithm. In: *IEEE Signal Process. Lett.* **7**(6), 164–168 (2000)
6. Yegnanarayana, B., Raghunathan, A.: Representation of images through group delay functions. In: *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-35**(2), 237–240 (1987)
7. Hahn, S.L.: *Hilbert Transforms in Signal Processing*. pp. 337–340. Art Tech House Inc., Norwood (1996)
8. Narasimhan, S.V., Veena, S.: *Signal Processing: Principles and Implementation*. Narosa Publishers/Alpha Science Publishers, New Delhi/UK (2008)